BRINT: Binary Rotation Invariant and Noise Tolerant Texture Classification

Li Liu, Yunli Long, Paul Fieguth, Songyang Lao, and Guoying Zhao

Abstract—In this paper we propose a simple, efficient, yet robust multi-resolution approach to texture classification — Binary Rotation Invariant and Noise Tolerant (BRINT). The proposed approach is very fast to build, very compact while remaining robust to illumination variations, rotation changes and noise. We develop a novel and simple strategy to compute a local binary descriptor based on the conventional LBP approach, preserving the advantageous characteristics of uniform LBP. Points are sampled in a circular neighborhood, but keeping the number of bins in a single-scale LBP histogram constant and small, such that arbitrarily large circular neighborhoods can be sampled and compactly encoded over a number of scales. There is no necessity to learn a texton dictionary, as in methods based on clustering, and no tuning of parameters is required to deal with different datasets.

Extensive experimental results on representative texture databases show that the proposed BRINT not only demonstrates superior performance to a number of recent state-of-the-art LBP variants under normal conditions but also performs significantly and consistently better in presence of noise due to its high distinctiveness and robustness. This noise robustness characteristic of the proposed BRINT is evaluated quantitatively with different artificially generated types and levels of noise (including Gaussian, salt and pepper and speckle noise) in natural texture images.

Index Terms—Texture descriptors, rotation invariance, local binary pattern (LBP), feature extraction, texture analysis

I. INTRODUCTION

Texture is a fundamental characteristic of the appearance of virtually all natural surfaces and is ubiquitous in natural images. Texture classification, as one of the major problems in texture analysis, has received considerable attention during the past decades due to its value both in understanding how the texture recognition process works in humans as well as in the important role it plays in the field of computer vision and pattern recognition [1]. Typical applications of texture classification include medical image analysis and understanding, object recognition, content-based image retrieval, remote sensing, industrial inspection, and document classification.

The texture classification problem is conventionally divided into the two subproblems. It is generally agreed that the extraction of powerful texture features is of more importance to the success of texture classification and, consequently, most research in texture classification focuses on the feature extraction part [1], with extensive surveys [1]. Nevertheless it remains a challenge to design texture features which are computationally efficient, highly discriminative and effective, robust to imaging environment changes (including changes in illumination, rotation, view point, scaling and occlusion) and insensitive to noise.

Recently, the Bag-of-Words (BoW) paradigm, representing texture images as histograms over a discrete vocabulary of local features, has proved effective in providing texture features [2]–[7]. Representing a texture image using the BoW model typically involves the following three steps:

(i) Local texture descriptors: extracting distinctive and robust texture features from local regions;

(ii) Texton dictionary formulation: generating a set of representative vectors (i.e., textons or dictionary atoms) learned from a large number of texture features;

(iii) Global statistical histogram computation: representing a texture images statistically as a compact histogram over the learned texton dictionary.

Within the BoW framework, the focus of attention has been on the design of local texture descriptors capable of achieving local invariance [2], [4]–[7]. These descriptors can be classified as dense or sparse, with the sparse approaches, such as SPIN, SIFT and RIFT [4], [10], requiring a process of detecting salient regions before applying the texture descriptors, leading to issues of implementation and computational complexity and instability. In contrast, dense approaches, applying texture descriptors pixel by pixel are more popular. Important dense textures descriptors include Gabor wavelets [8], LM filters [5], MR8 filters [5], BIF features [7], LBP [2], Patch descriptor [6] and RP random features [3] and many others [4].

Among local texture descriptors, LBP [2], [11] has emerged as one of the most prominent and has attracted increasing attention in the field of image processing and computer vision due to its outstanding advantages: (1) ease of implementation, (2) no need for pre-training, (3) invariance to monotonic illumination changes, and (4) low computational complexity, making LBP a preferred choice for many applications.
Although originally proposed for texture analysis, the LBP method has been successfully applied to many diverse areas of image processing: dynamic texture recognition, remote sensing, fingerprint matching, visual inspection, image retrieval, biomedical image analysis, face image analysis, motion analysis, edge detection, and environment modeling [12]–[17]. Consequently many LBP variants are present in the recent literature.

Although significant progress has been made, most LBP variants still have prominent limitations, mostly the sensitivity to noise [19], [21], and the limiting of LBP variants to three scales, failing to capture long range texture information [19], [21], [23]. Although some efforts have been made to include complementary filtering techniques [21], [24], these increase the computational complexity, running counter to the computational efficiency property of the LBP method.

In this paper, we propose a novel, computationally simple approach, the Binary Rotation Invariant and Noise Tolerant (BRINT) descriptor, which has the following outstanding advantages: It is highly discriminative, has low computational complexity, is highly robust to noise and rotation, and allows for compactly encoding a number of scales and arbitrarily large circular neighborhoods. At the feature extraction stage there is no pre-learning process and no additional parameters to be learned.

We derive a rotation invariant and noise tolerant local binary pattern descriptor, dubbed as BRINT \( S_{r,q} \), based on a circularly symmetric neighbor set of \( 8q \) members on a circle of radius \( r \). Parameter \( q \) controls the quantization of the angular space, and \( r \) determines the spatial scale of the BRINT \( S_{r,q} \) operator, which produces a histogram feature of constant dimensionality at any spatial scale \( r \) with arbitrary large number of sampling points \( 8q \) for each texture image.

Motivated by the recent CLBP approach, which was proposed by Guo et al. [25] to include both the signs and the magnitudes components between a given central pixel and its neighbors and the center pixel intensity in order to improve the discriminative power of the original LBP operator, we extend BRINT to include a magnitude component and to code the intensity of the center pixel. Based on these methods we develop a discriminative and robust combination for multiresolution analysis, which will be demonstrated experimentally to perform robustly against changes in gray-scale, rotation, and noise.

The remainder of this paper is organized as follows. A brief review of LBP and CLBP is given in Section II. Section III presents the motivation and the development of the new proposed BRINT approach in detail, as well as the multiresolution analysis and a brief overview of the classification process. Comprehensive experimental results and comparative evaluation are given in Section IV. Section V concludes the paper. A preliminary version of this work appeared in [9].

II. LBP AND CLBP

Despite the great success of LBP in computer vision and image processing, the original LBP descriptor [11] has some limitations: producing long histograms which are not rotation invariant; capturing only the very local texture structure and being unable to exploit long range information; limited discriminative capability based purely on local binarized differences; and and lacking noise robustness. On the basis of these issues, many LBP variations have been developed (see surveys [12], [13]), focusing on different aspects of the original LBP descriptor.

Dimensionality Reduction and Rotation Invariance

Most common is to reduce the feature length based on some rules, where influential work has been done by Ojala et al. [2] who proposed three important descriptors: rotation invariant LBP (LBP\(^\text{rot}\)), uniform LBP (LBP\(^\text{uniform}\)), and rotation invariant uniform LBP (LBP\(^\text{uniform}\text{rot}\)). Of these, LBP\(^\text{uniform}\text{rot}\), described in Section II-A, has become the most popular since it reduces the dimensionality of the original LBP significantly and achieves improved discriminative ability.

Discriminative Power

There are two approaches to improve discriminative power: reclassifying the original LBP patterns to form more discriminative clusters, or including other local binary descriptors. Noticeable examples include the Hamming LBP [26], which regroups nonuniform patterns based on Hamming distance instead of collecting them into a single bin as LBP\(^\text{uniform}\text{rot}\) does, the CLBP approach [25] which is discussed in Section II-B, and the Extended LBP approach [27] which considers the local binary descriptors computed from local intensities, radial differences and angular differences.

Noise Robustness

Ahonen et al. introduced Soft LBP (SLBP) method [28] which allows multiple local binary patterns to be generated at each pixel position, to make the traditional LBP approach more robust to noise; however, SLBP is computationally expensive and is no longer strictly invariant to monotonic illumination changes. Tan and Triggs [29] introduced local ternary patterns (LTP), where the binary LBP code ia replaced by a ternary LTP code. The LTP method is more resistant to noise, but no longer strictly invariant to gray-scale changes. Liao et al. [21] proposed to use dominant LBP (DLBP) patterns which considers the most frequently occurred patterns in a texture image. The Median Binary Pattern (MBP) proposed in [30] claims increased robustness to impulse noise such as salt-and-pepper noise, but MBP was only explored in a local 3\( \times \)3-patch. Fathi et al. [18] proposed a noise tolerant method based on the traditional LBP by combining a circular majority voting filter and a new LBP variant which regroups the nonuniform LBP patterns in order to gain more discriminability. Rajia et al. [22] proposed Optimized Local Ternary Patterns (OLTP) based on LTP in order to reduce feature dimensionality, however the authors did not extend OLTP to multiscale analysis. Ren et al. [20] proposed a much more efficient Noise Resistant Local Binary Pattern (NRLBP) approach based on the SLBP method, but it is computationally expensive to generalize to larger scales with a bigger number neighboring points.

Combining with Other Approaches

Ojala et al. [2] proposed a local contrast descriptor VAR to combine with LBP; It was recommended in [21] that Gabor filters and LBP-based features are mutually complementary.

---

1 A comprehensive bibliography of LBP methodology can be found at http://www.cse.oulu.fi/MVG/LBP_Bibliography/.
because LBP captures the local texture structure, whereas Gabor filters extract global texture information. Ahonen et al. proposed an approach named LBP histogram Fourier features (LBP-HF) [24], which combines the LBP and the discrete Fourier transform (DFT). Khellah [19] introduced a Dominant Neighborhood Structure (DNS) method which extracts global rotation-invariant features from the detected image dominant neighborhood structure to complement LBP.

A. Local Binary Patterns (LBP)

The original LBP method, proposed by Ojala et al. [11] in 1996, characterizes the spatial structure of a local image texture by thresholding a \( 3 \times 3 \) square neighborhood with the value of the center pixel and considering only the sign information to form a local binary pattern. A more general formulation defined on circular symmetric neighborhood systems was proposed in [2] that allowed for multi-resolution analysis and rotation invariance. Formally, given a pixel \( x_c \) in the image, the LBP pattern is computed by comparing its value with those of its \( p \) neighboring pixels

\[
\Sigma_{r,p} = [x_{r,p,0}, \ldots, x_{r,p,p-1}]^T
\]

that are evenly distributed in angle on a circle of radius \( r \) centered at center \( x_c \), as in Fig. 1, such that the LBP response is calculated as

\[
\text{LBP}_{r,p} = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_c)2^n, \quad s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}
\]

(1)

where \( s() \) is the sign function. Relative to the origin at \( (0,0) \) of the center pixel \( x_c \), the coordinates of the neighbors are given by \(-r \sin(2\pi n/p), r \cos(2\pi n/p)\). The gray values of neighbors which do not fall exactly in the center of pixels are estimated by interpolation.

Given an \( N \times M \) texture image \( I \), a LBP pattern \( \text{LBP}_{r,p}(i,j) \) can be computed at each pixel \((i,j)\). A texture image can be characterized by the probability distribution of the LBP patterns. Formally, the whole textured image \( I \) is represented by a LBP histogram vector \( \mathbf{h} \):

\[
\mathbf{h}(k) = \sum_{i=1}^{N} \sum_{j=1}^{M} \delta(\text{LBP}_{r,p}(i,j) - k)
\]

(2)

where \( 0 \leq k < 2^p \) is the number of LBP patterns. To be able to include textural information at different scales, the LBP operator was later extended to use neighborhoods of different sizes [2], with values of \((r,p)\) selected as \((1,8),(2,16),(3,24),\ldots,(r,8r)\).

A rotation invariant version \( \text{LBP}_{r,p}^{ri} \) of the original LBP descriptor was proposed by Pietikäinen et al. in [34]. The \( \text{LBP}_{r,p}^{ri} \) descriptor uses only the rotation invariant LBP patterns

\[
\text{LBP}_{r,p}^{ri} = \min\{\text{ROD}(\text{LBP}_{r,p},i) \mid i = 0, 1, \ldots, p - 1\}
\]

(3)

where \( \text{ROD}(x,i) \) performs a circular \( i \)-step bit-wise right shift on \( x \) \( i \) times. Keeping only those rotationally-unique patterns leads to a significant reduction in feature dimensionality, as shown in Table I, although beyond one scale the number of bins remains large. The \( \text{LBP}_{r,p}^{ri} \) descriptor was found to have poor performance [2], [34], therefore it has received little attention.

In order to obtain improved rotation invariance and to further reduce the dimensionality of the LBP histogram feature, building on LBP\( r,p \) Ojala et al. [2] proposed the “rotation invariant uniform” patterns \( \text{LBP}_{r,p}^{ru} \), the collection of those rotation invariance patterns having a \( U \) value of at most 2:

\[
\text{LBP}_{r,p}^{ru} = \left\{ \sum_{n=0}^{p-1} s(x_{r,p,n} - x_c), \quad \text{if } U(\text{LBP}_{r,p}) \leq 2 \right\}
\]

(4)

where

\[
U(\text{LBP}_{r,p}) = \sum_{n=0}^{p-1} |s(x_{r,p,n} - x_c) - s(x_{r,p,n+1} - x_c)|.
\]

(5)

There are \( p + 1 \) distinct groups of rotation invariant uniform patterns, with the rest considered as “nonuniform” patterns which are merged into one group, leading to a much lower dimensional histogram representation for the whole image, as shown in Table I. The success of the \( \text{LBP}_{r,p}^{ru} \) operator comes from the experimental observation that the uniform patterns appear to be fundamental properties of local image textures [2], representing salient local texture structure.

B. Completed Local Binary Patterns (CLBP)

Completed Local Binary Patterns (CLBP) [25] consist of three LBP-like descriptors: \( \text{CLBP}_C \), \( \text{CLBP}_S \) and \( \text{CLBP}_M \) which include information on the center pixel, signed differences, and magnitudes of differences, respectively, with the variants tested to improve the discriminative power of the original LBP operator. The \( \text{CLBP}_S \) descriptor is exactly the same as the original LBP descriptor; \( \text{CLBP}_C \) thresholds the
central pixel against the global mean gray value of the whole image:

$$\text{CLBP}_C = s \left( x_c - \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} I(i,j) \right). \quad (6)$$

CLBP_M performs a binary comparison between the absolute value of the difference between the central pixel and its neighbors and a global threshold to generate an LBP-like code:

$$\text{CLBP}_{M,r;p} = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_c) - \mu_{r,p}^n 2^n \quad (7)$$

where the global threshold $\mu_{r,p}^n$ used by Guo et al. [25] is computed as:

$$\mu_{r,p}^n = \frac{\sum_{j=r+1}^{N-r} \sum_{i=r+1}^{M-r} x_{r,p,n}(i,j) - x(i,j)}{(M-2r)(N-2r)p} \quad (8)$$

Since the CLBP approach adopts the ‘uniform and rotation invariant’ scheme for texture representation, clearly it inherits the main characteristics of the traditional LBP$_{r;p}$ (i.e. CLBP$_{S_{r;p}}$) descriptor. Moreover, due to the combination of three complementary descriptors CLBP$_{M_{r;p}}$, CLBP$_C$ and CLBP$_{S_{r;p}}$ jointly, CLBP has provided better texture classification performance than traditional LBP$_{r;p}$, but leads to much higher dimensionality.

III. BRINT: A Binary Rotation Invariant and Noise Tolerant Descriptor

A. Motivation

Although the original LBP approach is attractive for its conceptual simplicity and efficient computation, a straightforward application of the original LBP$_{r;p}$ histogram features is limited:

1) As shown in Table I, the original LBP operator produces rather long histograms ($2^p$ distinct values), overwhelmingly large even for small neighborhoods, leading to poor discriminant power and large storage requirements.

2) The LBP operator captures only the very local structure of the texture, appropriate for macro-textures but not for macro-textures. Because the LBP dimensionality becomes intractable as the sampling radius increases, it is difficult to collect information from a larger area.

3) The original LBP codes computed based on (1) are sensitive to image rotation.

4) LBP codes can be highly sensitive to noise: the slightest fluctuation above or below the value of the central pixel is treated the same way as a major contrast.

The rotation invariant descriptor LBP$_{r;p}^{riu}$ has received very limited attention, having shortcomings (1,2,4) listed above and in fact providing poor results for rotation invariant texture classification [34].

The LBP$_{r;p}^{riu}$ descriptor has avoided the disadvantages (1) and (2), which can be seen from Table I. However despite its clear advantages of dimensionality, gray scale and rotation invariance, and suitability for multi-resolution analysis, it suffers in terms of reliability and robustness as it only uses the uniform patterns and has minimal tolerance to noise.

The CLBP$_C$ + CLBP$_{S_{r;p}} +$ CLBP$_{M_{r;p}}$, abbreviated as CLBP_CSM, has been shown to perform better than LBP$_{r;p}$ ([25], due to the joint behavior of the three complementary LBP-like codes CLBP$_C$, CLBP$_S$ and CLBP$_M$, although this concatenation leads to a feature vector relatively high dimensionality (Table I). In standard CLBP_CSM applications, typically three scales are considered, with a corresponding dimensionality of 2200. The CLBP_CSM approach adopted in [35], utilizes five scales to extract texture feature, leading to an even higher dimensionality of 8040. For a multi-resolution analysis, with non-local features based on a larger number of scales, the increased dimensionality leads to challenges in storage and reliable classifier learning.

All of the discussed descriptors share one or more weaknesses of noise sensitivity, high dimensionality, and/or information insufficiency. Though all of the LBP-based approaches are computationally simple at the feature extraction step, except for LBP$_{r;p}^{riu}$, the other descriptors are all computationally expensive at the classification stage due to the high dimensionality of the histogram feature vector. The inherent difficulty in extracting suitable features for robust texture classification lies in balancing the three competing goals of discriminativeness, low computational requirements, and a robustness to noise. The goal of this paper was to build on the advantageous characteristics of LBP, developing an approach which achieves a better balance among these three competing requirements, in particular increasing robustness to noise. Our concern with the reduced approaches of LBP$_{r;p}^{riu}$ and CLBP_CSM lies with the use of only the uniform LBP patterns, which appear to lack texture discriminability. Instead, the LBP$_{r;p}^*$, although having large dimensionality, possesses meaningful texture features and strikes us as a more promising starting point.

B. BRINT: Proposed Approach

1) BRINT$_S$ descriptor: The construction of the local BRINT$_S$ descriptor is illustrated in Fig. 2. Similar to the sampling scheme in the original LBP approach, we sample pixels around a central pixel $x_c$, however on any circle of radius $r$ we restrict the number of points sampled to be a multiple of eight, thus $p = 8q$ for positive integer $q$. So the neighbors of $x_c$ sampled on radius $r$ are $x_{r,q} = [x_{r,8q,0}, \ldots, x_{r,8q,8q-1}]^T$.

In contrast to original LBP, we transform the neighbor vector $x_{r,q}$ by local averaging along an arc, $y_{r,q,i} = \frac{1}{q} \sum_{k=0}^{q-1} x_{r,8q,(qi+k)}$, $i = 0, \ldots, 7$, as illustrated in Fig. 2, such that the number of neighbors in $y_{r,q}$ is always eight.

Given $y_{r,q} = [y_{r,q,0}, \cdots, y_{r,q,7}]^T$, we can trivially compute a binary pattern with respect to the center pixel, as in LBP:

$$\text{BNT}_S_{r,q} = \sum_{n=0}^{7} s(y_{r,q,n} - x_c) 2^n \quad (10)$$

where BNT$_S$ represents “Binary Noise Tolerant Sign”. One can easily see that for any parameter pair $(r, q)$ there are
The Idea

\[ \sum_{r=1}^{3} X_{r,8q} \]

\[ \frac{1}{q} \sum_{q=1}^{q} y_{r,q} \to a(y_{r,q} - x_{c}) \]

BNT_S

Fig. 2. Illustration of the proposed BNT_S descriptor which is designed to derive the proposed BRINT descriptor. The definition of the BNT_S descriptor, and a 3-scale example illustrating the construction of the proposed BNT_S descriptor. This figure is better read in color. Rather than directly subtracting the gray value \( x_c \) of the center pixel from the precise gray value of each neighboring pixel \( x_{r,8q}, i = 0, \ldots, 8q - 1 \), the proposed approach introduces a novel idea – Average-Before-Quantization (ABQ) – first transforming the original neighborhood into a new one \( y_{r,q}, i = 0, \ldots, 7 \), and then thresholding \( y_{r,q}, i = 0, \ldots, 7 \) at the gray value of the center pixel to generate a binary pattern. See text for further details.

An example

<table>
<thead>
<tr>
<th>Image Pixels</th>
<th>Averaging</th>
<th>Binary Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Fig. 3. A motivational example for illustration of noise robustness. Middle: A 7 \( \times \) 7-pixels image and its zero mean additive Gaussian noise added version. The conventional LBP responses are shown on the left, in contrast to the BNT_S pattern on the right. The BNT_S approach shows greater consistency in the presence of noise.

\[ 2^8 = 256 \text{ BNT}_S_{r,q} \text{ binary patterns in total.} \]

Furthermore, the transformation from \( \sum_{r=1}^{3} X_{r,8q} \) to \( \sum_{r=1}^{3} y_{r,q} \) makes the pattern more robust to noise, as is illustrated in an example in Fig. 3.

As rotation invariance is one of our stated objectives, and to avoid the limitations [13, 19, 21] of uniform patterns, we follow the inspiration of LBP_{r,p}^q and grouping equal versions of binary representations under rotations, assigning code numbers to the resulting groups. Formally, then, BINT_N_{r,q} is defined as

\[ \text{BINT}_N_{r,q} = \min \{ ROR(\text{BINT}_N_{r,q}), i = 0, \ldots, 7 \} \]

where rotation function \( ROR(\bullet, \bullet) \) is as in (3), reducing the number of histogram bins, for one scale, from 256 to 36. The motivation, then, for fixing the number of points in \( y_{r,q} \) to a constant 8 was to limit the growth in histogram bins with scale.

In terms of parameter \( q \), which controls the number of neighbors being sampled and averaged, we illustrate two reasonable sampling schemes in Fig. 4. Scheme 1, employed in BRINT1, should be more robust to noise, due to having more neighbors to average, however it may cause over-smoothing relative to Scheme 2, employed in method BRINT2.

Fig. 5 validates the basic behavior of BRINT_N_{r,q} as a function of the number of scales by contrasting its classification performance with that of the conventional LBP_{r,p}^q descriptor. The classification results show a significant jump in classification performance on the three Outex databases, outperforming the best results reported by Ojala et al. [2].

In terms of computation cost, the proposed BRINT_N descriptor does not imply an increase in complexity over the traditional LBP_{r,p}^{q=2}. In particular, BRINT_N always deals with local binary patterns based on 8 points, whereas for LBP_{r,p}^{q=2} the mapping from LBP to LBP_{r,p}^{q=2} requires a large lookup table having \( 2^q \) elements.

2) BRINT_M descriptor: Motivated by the striking classification results achieved by BRINT_S and considering the better performance of the CLBP_CSM feature over the single feature LBP_{r,p}^{q=2} proposed by Guo et al. [25], we would like to further capitalize on the CLBP_M descriptor by proposing BRINT_M.
where $\mu_l$ is the local thresholding value. Note that the CLBP_M descriptor defined in (7) of [25] uses the global threshold $\mu_0$ of (8), whereas in the original LBP operator the thresholding value is the center pixel value, which clearly varies from pixel to pixel. Therefore, instead of using a constant global threshold, we propose to use a locally varying one:

$$
\mu_{l,r,q} = \frac{1}{8} \sum_{n=0}^{7} z_{r,q,n}.
$$

(15)

With BNT_M defined, BRINT_M is defined as

$$
BRINT_M_{r,q} = \min \{ ROR(BNT_M_{r,q,i}) | i = 0, \ldots, 7 \}. 
$$

(16)

Fig. 6 compares the results of the proposed BRINT_M with the comparable CLBP methods, with BRINT_M significantly outperforming.

Finally, consistent with CLBP, we also represent the center pixel in one of two bins:

$$
BRINT_C_{r} = s(x_c - \mu_{l,r})
$$

(17)

where $\mu_{l,r}$ is the mean of the whole image excluding boundary pixels:

$$
\mu_{l,r} = \frac{1}{(M-2r)(N-2r)} \sum_{i=r+1}^{M-r} \sum_{j=r+1}^{N-r} x(i,j).
$$

(18)

C. MultiResolution BRINT

The proposed BRINT descriptors were, so far, extracted from a single resolution with a circularly symmetric neighbor set of $8q$ pixels placed on a circle of radius $r$. Given that one goal of our approach is to cope with a large number of different scales, by altering $r$ we can create operators for different spatial resolutions, ideally representing a textured patch by concatenating binary histograms from multiple resolutions into a single histogram, as illustrated in Fig. 7, clearly requiring that the histogram feature produced at each resolution be of low dimension.

Since BRINT_CSM, the joint histogram of BRINT_C, BRINT_S and BRINT_M, has a very high dimensionality of $36*36*2 = 2592$, in order to reduce the number of bins needed we adopt the BRINT_CSM_M descriptor, meaning the joint histogram $BRINT_C * BRINT_S * BRINT_M$, producing a histogram of much lower dimensionality: $36 * 2 + 36 * 2 = 144$. As a point of comparison, in the experimental results we will also evaluate $BRINT_S * BRINT_M$ having a dimensionality of $36 + 36 = 72$.

D. Classification

The actual classification is performed via one of two popular classifiers:
1) The Nearest Neighbor Classifier (NNC) applied to the normalized BRINT histogram feature vectors $h_i$ and $h_j$, using the $\chi^2$ distance metric as in [3], [5], [6], [25], [38].

2) The nonlinear Support Vector Machine (SVM) of [43], where the benefits of SVMs for histogram-based classification have clearly been demonstrated in [4], [21], [31]. Kernels commonly used include polynomials, Gaussian Radial Basis Functions and exponential Chi-Square kernel. Motivated by [4], [21], [31], we focus on the exponential $\chi^2$ kernel

$$K(h_i, h_j) = \exp(-\gamma \chi^2(h_i, h_j)),$$

(19)

where only one parameter $\gamma$ needs to be optimized. We use the one-against-one technique, which trains a classifier for each possible pair of classes.

IV. EXPERIMENTAL EVALUATION

A. Image Data and Experimental Set up

For our experimental evaluation we have used six texture datasets, summarized in Table II, derived from the four most commonly used texture sources: the Brodatz album [32], the CUReT database [6] and KTH-TIPS2b [42]. The Brodatz database is perhaps the best known benchmark for evaluating texture classification algorithms. Performing classification on the entire database is challenging due to the relatively large number of texture classes, the small number of examples for each class, and the lack of intra-class variation.

1) Experiment # 1: There are 24 different homogeneous texture classes selected from the Outex texture database [33], with each class having only one sample of size $538 \times 746$-pixels. The 24 different texture samples are imaged under different lighting and rotations conditions. Three experimental test suites Outex_TC10, Outex_TC12_000 and Outex_TC12_001, summarized in Table II, were developed by Ojala et al. [2] as benchmark datasets for rotation and illumination invariant texture classification. For all the three test suites, the classifier is trained with 20 reference images of the 'inca' illumination condition and angle $0^\circ$, totaling 480 samples. The difference among these three test suites is in the testing set. For Outex_TC10, the remaining 3840 samples with 'inca' illumination, are used for testing the classifier. For Outex_TC12_000 and Outex_TC12_001, the classifier is tested with all 4320 images from fluorescent and sunlight lighting, respectively.

For the experiments on all three Outex databases, we first test the classification performance of the proposed approach on the original database and then assess the robustness of the proposed method under noisy conditions, where the original texture images are corrupted by zero-mean additive Gaussian noise with different Signal-to-Noise Ratios (SNRs) (defined as the ratio of signal power to the noise power). Moreover, we also test the classification performance of the proposed approach against impulse salt-and-pepper noise with different noise density ratio and multiplicative noise with zero mean and different variances, which is randomly and independently added to each image.

2) Experiment # 2: Brodatz was chosen to allow a direct comparison with the state-of-the-art results from [21]. There are 24 homogeneous texture classes\(^2\). Each image was partitioned into 25 nonoverlapping sub-images of size of $128 \times 128$, each of these downsampled to $64 \times 64$. 13 samples per class were selected randomly for training and the remaining 12 for testing.

For the CUReT database, we use the same subset of images which has been previously used in [3], [5], [6], [19], [25], [35], [38]: 61 texture classes each with 92 images under varying illumination direction but at constant scale. 46 samples per class were selected randomly for training and the remaining 46 for testing. It has been argued [5], [6], [39] that this scale constancy is a major drawback of CUReT, leading to KTH-TIPS2b [39], [42], with 3 viewing angles, 4 illuminants, and 9 different scales. We follow the training and testing scheme of [39]: training on three samples and testing on unseen samples.

For Brodatz and CUReT, results for texture classification under random Gaussian noisy environment are also provided. Training and testing scheme is the same as in noise-free situation.

B. Methods in Comparison and Implementation Details

We will be performing a comparative evaluation of our proposed approach, where the abbreviations of the proposed

---

### TABLE II

<table>
<thead>
<tr>
<th>Texture Dataset</th>
<th>Rotation</th>
<th>Illumination Variation</th>
<th>Scale Variation</th>
<th>Texture Class</th>
<th>Sample per Class</th>
<th>Training Samples</th>
<th>Test Samples</th>
<th>Total Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outex_TC10</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>24</td>
<td>128 × 128</td>
<td>181</td>
<td>40</td>
<td>221</td>
</tr>
<tr>
<td>Outex_TC12_000</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>24</td>
<td>128 × 128</td>
<td>181</td>
<td>40</td>
<td>221</td>
</tr>
<tr>
<td>Outex_TC12_001</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>24</td>
<td>128 × 128</td>
<td>181</td>
<td>40</td>
<td>221</td>
</tr>
</tbody>
</table>

### TABLE III

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRINT_S</td>
<td>Binary rotation invariant and noise tolerant descriptor based on sign component BRINT_rsaq</td>
</tr>
<tr>
<td>BRINT_M</td>
<td>Binary rotation invariant and noise tolerant descriptor based on magnitude component BRINT_rmag</td>
</tr>
<tr>
<td>BRINT_C (CLBP_C)</td>
<td>Binary pattern for the center pixel</td>
</tr>
<tr>
<td>BRINT_S_M</td>
<td>Concatenation of BRINT_S and BRINT_M</td>
</tr>
<tr>
<td>BRINT_CS (CLBP_M)</td>
<td>Joint distribution of BRINT_C and BRINT_S</td>
</tr>
<tr>
<td>BRINT_CM (CLBP_M)</td>
<td>Joint distribution of BRINT_C and BRINT_M</td>
</tr>
<tr>
<td>BRINT_CS_CM</td>
<td>Concatenation of BRINT_CS and BRINT_CM</td>
</tr>
<tr>
<td>LBP + (CLBP_S_r,p)</td>
<td>Rotation invariant LBP</td>
</tr>
<tr>
<td>LTP + (CLBP_S_r,p)</td>
<td>Rotation invariant uniform LTP</td>
</tr>
<tr>
<td>CLBP_M + (CLBP_S_r,p)</td>
<td>Rotation invariant magnitude LBP</td>
</tr>
<tr>
<td>CLBP_M + (CLBP_M_r,p)</td>
<td>Rotation invariant uniform magnitude LBP</td>
</tr>
<tr>
<td>CLBP_S + (CLBP_S_r,p)</td>
<td>Concatenation of BRINT_S and CLBP_M</td>
</tr>
<tr>
<td>CLBP_M + (CLBP_M_r,p)</td>
<td>Concatenation of BRINT_M and CLBP_M</td>
</tr>
<tr>
<td>CLBP_S + (CLBP_M_r,p)</td>
<td>Concatenation of BRINT_M and CLBP_M</td>
</tr>
<tr>
<td>CLBP_M + (CLBP_M_r,p)</td>
<td>Concatenation of BRINT_M and CLBP_M</td>
</tr>
</tbody>
</table>

---

\(^2\)The 24 Brodatz textures are D1, D4, D16, D19, D21, D24, D28, D32, D53, D54, D57, D65, D68, D77, D82, D84, D92, D93, D95, D98, D101, D102, D106, D111.
descriptor and state-of-the-art approaches are given in Table III:

1) \( \text{CLBP}_{\text{CS}}^{ri} \), \( \text{CM}^{ri} \): The rotation invariant CLBP approach parallel to our proposed BRINT_{CS} CM feature.

2) \( \text{CLBP}_{\text{CS}}^{ri} \text{CM}_{r;p}^{ri} \): The rotation invariant and uniform CLBP method parallel to our proposed BRINT_{CS} CM feature.

3) \( \text{DLBP+NGF} \) [21]: The fused features of the DLBP features and the normalized Gabor filter response average magnitudes (NGF). It is worth mentioning that the DLBP approach needs pretraining and the dimensionality of the DLBP feature varies with the training image.

4) \( \text{LTP} \) [29]: The recommended \( \text{LTF}_{r;p}^{ri} \) is used. Here we implemented a nine scale descriptor, where the associated parameter settings can be seen in Table IV.

5) \( \text{CLBP} \) [25]: The recommended fused descriptor \( \text{CLBP}_{\text{CSM}} \) (i.e. \( \text{CLBP}_{\text{CS}}^{ri2} \text{CM}_{r;p}^{ri2} \)) is used, however only a 3-scale \( \text{CLBP}_{\text{CSM}} \) is implemented due to the high dimensionality limitation mentioned in Table I.

6) \( \text{LBP} \) [2]: The traditional rotation invariant uniform feature proposed by Ojala et al. [2]. We use a 3-scale descriptor as recommended by the authors.

7) \( \text{DNS+LBP} \) [19]: The fused feature of Dominant Neighborhood Structure approach and the conventional LBP approach proposed by Khellah [19] claimed to have noise robustness.

8) \( \text{disCLBP} \) [15]: The discriminative descriptor obtained by a learning framework proposed by Guo et al. [15]. Due to the high dimensionality of the descriptor at larger scales, we use a 3-scale descriptor \( \text{dis}(S+M)_{r;p}^{ri2} \) as recommended by the authors.

9) \( \text{LBP}_{r;p;k}^{NT} \) [18]: A circular majority voting filter to achieve noise robustness, followed by a scheme to regroup the nonuniform LBP patterns into several different classes instead of classifying them into a single class as \( \text{LBP}_{r;p}^{ri2} \). Parameter \( k \) acts as the size of kernel in the circular majority voting filter, controlling the number of noisy bits that should be filtered in the obtained LBP pattern. As suggested by Fathi et al. [18], parameter \( k \) is set as 1, 3 and 4 for \( p = 8, 16 \) and 24 respectively. We implemented a multiresolution (nine scales) \( \text{LBP}_{r;p;k}^{NT} \) (MS9), however Fathi et al. [18] only considered three scales in their work.

10) \( \text{NRLBP} \) [20]: We implemented a multiresolution \( \text{NRLBP}_{r;p}^{ri2} \) descriptor: \( \text{NRLBP}_{r;p}^{ri2} r = 1, \ldots, 9 \), though Ren et al. [20] only evaluated the first scale \( (r, p) = (1, 8) \) in their original paper. The reason that the number of neighboring points \( p \) is kept 8 for each radius \( r \) is because the extraction of the NRLBP feature requires building up a lookup table of size \( 3p \) which is extremely expensive in terms of both computation time and memory cost.

Each texture sample is preprocessed: normalized to zero mean and unit standard deviation. For the CUReT and Brodatz databases, all results are reported over 100 random partitions of training and testing sets. For SVM classification, we use the publicly available LibSVM library [41]. The parameters \( C \) and \( \gamma \) are searched exponentially in the ranges of \( \left[2^{-5}, 2^{18}\right] \) and \( \left[2^{-15}, 2^8\right] \), respectively, with a step size of \( 2^1 \) to probe the highest classification rate. However, in our experiments setting \( C = 10^6 \) and \( \gamma = 0.01 \) give very good performance. In the additive Gaussian noise environment, the SNRs tested here are 100, 30, 15, 10, 5 and 3, corresponding to 20db, 14.74db, 11.76db, 10db, 7db and 4.77db respectively. The noise density ratios of the salt-and-pepper noise tested are \( \rho = 5\%, 10\%, 20\%, 30\%, 40\% \). The multiplicative noise tested is with zero mean and different variances \( \nu = 0.02, 0.05, 0.1, 0.15, 0.2, 0.3 \).

C. Results for Experiment # 1

Fig. 8 plots the classification performance of different BRINT combination schemes as a function of number of scales. There is a trend of increasing classification performance as the number of scales increases. It is apparent that the BRINT_{CS} CM feature performs the best, therefore the BRINT_{CS} CM descriptor will be our proposed choice and will be further evaluated.

Fig. 9 compares the two sampling schemes for the proposed approach, using the OuteX_TC12_000 database. Here we can see that sampling scheme 2 produced better classification performance than sampling scheme 1, believed to be because
Fig. 8. Classification rates as a function of number of scales, with the same experimental setup as in Fig. 5, using a NNC classifier. Of the combinations tried, BRINT2_CS_CM performs the best.

Fig. 9. Comparing the classification performance of the two sampling schemes of Fig. 4 on Outex_TC12_000. The experimental setup is the same as in Fig. 8. Scheme 2 performs better and will be adopted.

sampling scheme 1 oversmooths the local texture structure, resulting in lost texture information.

Table V compares the classification performance of the proposed BRINT2_CS_CM descriptor with those of CLBP [25] on the three Outex databases. We observe that BRINT2 performs significantly and consistently better than both $ri$ and $riu2$ forms of CLBP, both in single-scale and multiple-scale cases. The striking performance of BRINT2_CS_CM clearly demonstrates that the concatenated marginal distributions of the proposed basic BRINT_C, BRINT_S and BRINT_M codes and the novel “averaging before binarization” scheme turns out to be a very powerful representation of image texture. The use of multiple scales offers significant improvements over single-scale analysis, consistent with earlier results in Figs. 8 and 9, showing that the approach is making effective use of interactions between the center pixel and more distant pixels. To the best of our knowledge, the proposed approach produced classification scores which we believe to be the best reported for Outex_TC12_000 and Outex_TC12_001. Keeping in mind the variations in illumination and rotation present in the Outex databases, the results in Table V firmly demonstrate the illumination and rotation invariance property of the proposed BRINT_CS_CM approach.

Table VI compares the best classification scores achieved by the proposed BRINT2_CS_CM method using nine scales (MS9) in comparison with state-of-the-art texture classification methods on all three Outex test suites. Despite not being
customized to the separate test suites, our multi-scale BRINT2 descriptor produces what we believe to be the best reported results on all three suites, regardless whether NNC or SVM is used. We would also point out that except for the proposed BRINT, CLBP_CSM [25] and CLBC_CSM [35] approaches, the remaining descriptors listed in Table VI require an extra learning process to obtain the lexicon dictionary, requiring additional parameters or computational burden.

The preceding discussion allows us to assert that the proposed multi-scale BRINT2 approach outperforms the conventional multi-scale CLBP approach on the Outex test suites. We now wish to examine the robustness of our method against noise to test applicability to real-world applications, thus the original texture images from Experiment #1 have been subject to added Gaussian noise.

Table VII quite clearly shows the noise-robustness offered by the BRINT approach: similar classification rates are seen in the near-absence of noise (SNR=100), however the de-
single scale is clearly seen, as is the significant performance of BRINT2 and conventional CLBP descriptors under high noise found. Clearly, it can be observed from Table VIII that the results are given in Table VIII, where the proposed BRINT and those by the state-of-the-art methods on the Outex_TC12_000 test suite injected with additive Gaussian noise (corresponding to the results on the Outex_TC12_000 shown in Table VII). The √ mark indicates statistically significant existence. The bracketed values are the McNemar χ² square statistics and the p values (p<0.000, p<0.004).

Finally, Table IX and Table X compare the classification accuracy highlighted in bold. The NNC classifier is used.

Fig. 10. A comparison of classification performance under severe noise (SNR=5), both (a) as a function of the single scale used, and (b) as a function of the number of scales. The strength of BRINT2 over CLBP is clear, as is the benefit of forming features over as many scales as possible.

D. Results for Experiment # 2

The classification results on the original Brodatz databases are listed in Table XI. The proposed BRINT1 method with a NNC classifier performs the best at 100% accuracy, however honestly all of the tested methods achieve very high classification accuracies here, since all 24 tested textures are relatively homogeneous and have small intra-class variations caused by rotation and illumination variations, a relatively easy problem for classification.

Instead, the noise-corrupted Brodatz database is expected to introduce greater challenges, with results listed in Table XII. We specifically compare with DLBP+NGF [21], which is one
TABLE XII

<table>
<thead>
<tr>
<th>Features</th>
<th>Classification Accuracies (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NNC [85]</td>
</tr>
<tr>
<td>BRINT1_S_CM (MSV)</td>
<td>98.09</td>
</tr>
<tr>
<td>BRINT1_CS_CM (MSV)</td>
<td>98.92</td>
</tr>
<tr>
<td>BRINT2_S_CM (MSV)</td>
<td>96.86</td>
</tr>
<tr>
<td>BRINT2_CS_CM (MSV)</td>
<td>96.86</td>
</tr>
</tbody>
</table>

TABLE XIII

<table>
<thead>
<tr>
<th>Curves</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNC</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td></td>
</tr>
<tr>
<td>LBPv2+NGF [19]</td>
<td>[19]</td>
</tr>
<tr>
<td>CLBP [85]</td>
<td>[85]</td>
</tr>
<tr>
<td>VZ-ART [82]</td>
<td>[82]</td>
</tr>
<tr>
<td>SVM</td>
<td></td>
</tr>
</tbody>
</table>

TABLE XIV

<table>
<thead>
<tr>
<th>Curves</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNC</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td></td>
</tr>
<tr>
<td>LBPv2+NGF [19]</td>
<td>[19]</td>
</tr>
<tr>
<td>CLBP [85]</td>
<td>[85]</td>
</tr>
<tr>
<td>VZ-ART [82]</td>
<td>[82]</td>
</tr>
<tr>
<td>SVM</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 11. Classification performance of the proposed approach with various state-of-the-art results on the KTHTIPS2b texture database. The BRINT results are based on nine scales and NNC. All results are computed by us, except for those of MWLD and SIFT, which are quoted from [40].

V. CONCLUSIONS

The multi-resolution LBP$_{v2}$ and the more recent LBP$_{v2}$ disparities descriptors have been proved to be two powerful measures of image texture [2], [25]. However, they have also been shown to have serious limitations including the instability of the uniform patterns, the lack of noise robustness, the inability to encode a large number of different local neighborhoods, an incapability to cope with large local neighborhoods, and high dimensionality (CLBP) [13], [21], [23]. In order to avoid these problems, we have presented BRINT, a theoretically and computationally simple, noise tolerant yet highly effective multi-resolution descriptor for rotation invariant texture classification. The proposed BRINT

Table XV confirms the noise robustness of the proposed BRINT approach, emphasizing that no smoothing is necessary. The absence of spatial smoothing is a significant advantage for BRINT, as local spatial information is important for texture classification, whereas pre-smoothing can suppress important local texture information, a serious drawback for texture recognition in low-noise situations.
TABLE XV

Classification accuracies (%) on the noisy-cooperated-OUTEX database, comparing the methods with or without PRE-GAUSSIAN smoothing. All results are reported over 50 random partitionings of the training and test sets. For each test, the highest mean classification accuracies are highlighted in bold. GAUSSIAN SMOOTHING FILTER with σ = 1.5 is used. For classification, the SVM classifier is used.

<table>
<thead>
<tr>
<th>Method</th>
<th>Classification accuracies (%)</th>
<th>SNR=15</th>
<th>SNR=10</th>
<th>SNR=30</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE-GAUSSIAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRINT_C2M (Ms)</td>
<td>94.0 ± 1.9</td>
<td>93.8</td>
<td>92.6</td>
<td>90.1</td>
</tr>
<tr>
<td>BRINT_LTP (Ms)</td>
<td>92.9 ± 2.1</td>
<td>91.0</td>
<td>88.5</td>
<td>84.9</td>
</tr>
<tr>
<td>BRINT_LTP (Ms)</td>
<td>91.5 ± 2.3</td>
<td>90.4</td>
<td>87.2</td>
<td>84.1</td>
</tr>
</tbody>
</table>

The robustness of the proposed approach to image rotation and noise has been validated with extensive experiments on six different texture datasets. This noise robustness characteristic is evaluated quantitatively with different artificially generated types and levels of noise (including Gaussian, salt and pepper and multiplicative noise) in natural texture images. The proposed approach to produce consistently good classification results on all of the datasets, most significantly outperforming the state-of-the-art methods in high noise conditions.

The current work has focused on texture classification. Future interest lies how to exploit the proposed descriptor for the domain of face recognition and object recognition.

REFERENCES

Li Liu received the B.S. degree in communication engineering, the M.S. degree in photogrammetry and remote sensing and the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2003, 2005 and 2012, respectively. He is currently a post doctoral fellow in the School of Information System and Management. She was a Visiting Student at the University of Waterloo, Canada, from 2008 to 2010. She has held a visiting appointment at the Beijing University in China. Dr. Liu is a Co-Chair of the International Workshop on Robust local descriptors for computer vision (RoLoD) at ACCV2014. Her current research interests include computer vision, texture analysis, pattern recognition and video analysis and retrieval.

Yunli Long received the M.S. degree in electronic science and technology and the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2005 and 2012, respectively. He is currently a post doctoral fellow in the School of Electronic Science and Engineering at the National University of Defense Technology. His current research interests include object detection and recognition, object tracking and video surveillance.

Songyang Lao received the B.S. degree in information system engineering and the Ph.D. degree in system engineering from the National University of Defense Technology, Changsha, China, in 1990 and 1996, respectively. He joined the faculty at the National University of Defense Technology in 1996, where he is currently a professor in School of Information System and Management. She was a Visiting Scholar with the Dublin City University, Irish, from 2004 to 2005. His current research interests include image processing and video analysis and human-computer interaction.

Guoying Zhao received the Ph.D. degree in computer science from the Chinese Academy of Sciences, Beijing, China, in 2005. She is currently an Associate Professor with the Center for Machine Vision Research, University of Oulu, Finland, where she has been a researcher since 2005. In 2011, she was selected to the highly competitive Academy Research Fellow position. She has authored or co-authored more than 100 papers in journals and conferences, and has served as a reviewer for many journals and conferences. She has lectured tutorials at ACCV, CVPR, FG, BMVC, CRV, AVSS. She is a member of program committees for many conferences, e.g., ICCV, CVPR, FG, BMVC, CRV, AVSS. She is IEEE Senior Member, and Editorial Board Member of International Journal of Applied Pattern Recognition and International Scholarly Research Network. Her current research interests include image and video descriptors, gait analysis, dynamic-texture recognition, facial-expression recognition, human motion analysis, and person identification.

Paul W. Fieguth (S’87 - M’96) received the B.A.Sc. degree from the University of Waterloo, Ontario, Canada, in 1991 and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1995, both degrees in electrical engineering. He joined the faculty at the University of Waterloo in 1996, where he is currently Professor in Systems Design Engineering. He has held visiting appointments at the University of Heidelberg in Germany, at INRIA/Sophia in France, at the Cambridge Research Laboratory in Boston, at Oxford University and the Rutherford Appleton Laboratory in England, and with postdoctoral positions in Computer Science at the University of Toronto and in Information and Decision Systems at MIT. His research interests include statistical signal and image processing, hierarchical algorithms, data fusion, and the interdisciplinary applications of such methods, particularly to remote sensing.